基于机器学习的重离子碰撞中QCD相变的研究

李甫鹏[1] 庞龙刚[1] 王新年[2]
1（华中师范大学 夸克与轻子教育部重点实验室 武汉 430079）
2（美国劳伦斯伯克利国家实验室 核科学部 伯克利 加州94720）

摘要 高能重离子碰撞将夸克和胶子从原子核中释放出来，形成一种新的核物质形态，即解禁闭的夸克胶子等离子体（Quark-Gluon Plasma，QGP）。研究普通核物质或强子共振气体到夸克胶子等离子体的相变，需要在超级计算机上数值求解格点量子色动力学（Quantum Chromodynamics，QCD）。但是，格点QCD只能给出零重子化学势以及附近可泰勒展开区域的核物质状态方程，并预测这种条件下QGP到强子共振气体之间的相变为平滑过渡。在不能做泰勒展开的有限重子化学势区域，格点QCD会遭遇著名的符号问题，无法给出有效的核物质状态方程以及QCD相变类型。本文综述了利用机器学习在核物质状态方程、相变分类以及临界点寻找方面的研究。这些研究大致分为两类：第一类在高能重离子碰撞实验数据以及相对论流体力学模拟和分子动力学模拟中，利用核碰撞末态粒子分布来确定核物质状态方程以及相变种类；另一类是利用机器学习直接帮助格点QCD的采样，解决有限密系统中的符号问题。

关键词 重离子碰撞，机器学习，QCD相变，深度神经网络，Ai4Science
中图分类号 TL99，O571.6，O572.2
DOI: 10.11889/j.0253-3219.2023.hjs.46.040014

Application of machine learning to the study of QCD transition in heavy ion collisions

LI Fupeng[1] PANG Longgang[1] WANG Xinnian[2]
1(*Key Laboratory of Quark and Lepton Physics* (*MOE*) *& Institute of Particle Physics*, *Central China Normal University*, *Wuhan 430079*, *China*)
2(*Nuclear Science Division*, *Lawrence Berkeley National Laboratory*, *Berkeley*, *CA94720*, *USA*)

Abstract In high-energy heavy ion collisions, quarks and gluons are released from the colliding nucleus to form a new state of nuclear matter called deconfined quark gluon plasma (QGP). To study the transition from normal nuclear matter or hadron resonance gas to QGP, non-perturbative quantum chromodynamics (QCD) must be solved on supercomputers using the lattice numerical method (lattice Quantum Chromodynamics, lattice QCD). However, lattice QCD only works for zero and small baryon chemical potential regions that can be described by the Taylor expansion and provides the nuclear equation of state (EoS) and QCD transition in these regions. For large baryon chemical potential regions that cannot be described by the Taylor expansion, lattice QCD fails to provide the nuclear

国家自然科学基金面上项目(No.1207509)、国家重点研发计划项目(No.2020YFE0202002)、广东基础和应用基础科学重大研究项目(No.2020B0301030008)、美国能源部(No.DE-AC02-05CH11231)资助
第一作者：李甫鹏，男，1995年出生，2021年于浙江工业大学获硕士学位，现为博士研究生，研究领域为理论物理
通信作者：庞龙刚，E-mail：lgpang@ccnu.edu.cn；王新年，E-mail：xnw@lbl.gov
收稿日期：2023-03-06，修回日期：2023-04-03
Supported by National Natural Science Foundation of China (No.1207509), the National Key Research and Development Program of China (No. 2020YFE0202002), the Guangdong Major Project of Basic and Applied Basic Research (No.2020B0301030008), the U.S.Department of Energy (No. DE-AC02-05CH11231)
First author: LI Fupeng, male, born in 1995, graduated from Zhejiang University of Technology with a master's degree in 2021, doctoral student, focusing on theoretical physics
Corresponding author: PANG Longgang, E-mail: lgpang@ccnu.edu.cn; WANG Xinnian, E-mail: xnw@lbl.gov
Received date: 2023-03-06, revised date: 2023-04-03

EoS and QCD transition owing to the famous sign problem. Machine learning helps to study the nuclear EoS and QCD phase transition. First, machine learning can determine the nuclear EoS and QCD transition using the momentum distribution of final state hadrons in heavy-ion collisions, with data from both heavy-ion collision experiments and relativistic hydrodynamic simulations. Second, it can contribute to the direct solution of the sign problem in lattice QCD. The present paper reviews the applications of machine learning to the study of the QCD phase transition in heavy-ion collisions. This study (1) introduces nuclear EoS and QCD transition as well as the difficulty of the lattice QCD method, (2) analyzes the nuclear EoS using Bayesian analysis, (3) identifies the nuclear EoS and QCD phase transition using different types of deep neural networks (e.g., convolutional neural network, point cloud network, and many-event averaging), (4) searches for critical self-similarity using a dynamical edge convolution-based graph neural network, (5) learns the quasi-particle mass using a physically informed network and auto-differentiation, (6) discards unphysical regions in the nuclear EoS with a critical endpoint using active learning, (7) discusses unsupervised learning for the nuclear liquid-gas phase transition, (8) determines the nuclear symmetry energy in heavy-ion collisions, (9) investigates Mach cones using deep learning assisted jet tomography, and (10) accelerates the sampling of lattice QCD configurations using a physically constrained neural network while solving the sign problem in lattice QCD using deep learning.

Key words　Heavy ion collisions, Machine learning, QCD phase transition, Deep neural network, Ai4Science

夸克和胶子总是以色单态的形式囚禁在强子中。“核子重如牛，对撞生新态”，李政道先生提出使用高能重离子碰撞产生高温高密的核物质环境，将夸克和胶子从原子核中释放出来，生成夸克胶子等离子体（Quark-Gluon Plasma，QGP）并研究核物质的解禁闭过程。高能重离子碰撞的实验和理论已经取得了长足发展。实验上已经发现，QGP存在的多个信号，比如高能核碰撞末态强子在动量空间存在很强的集体流，奇异夸克的增强，高能部分子喷注穿过QGP时的巨大能量损失。理论方面也建立了色玻璃凝聚，相对论流体力学，高能部分子输运的线性玻尔兹曼模型以及强子输运的相对论分子动力学模型，能够很好地描述高能部分子的能量损失以及中低横动量区末态强子的动量分布。使用实验数据与理论模型的对比，研究者发现QGP是实验室中产生的温度最高，磁场最强，剪切黏滞系数最小，旋转速度最快的流体，又称核液滴。关于相对论重离子碰撞更加详尽的信息，请参考文献[1−3]。

在高能重离子碰撞理论模型与实验数据的对比过程中，研究者遇到极大的困难。高能核碰撞是典型的反问题，QGP的寿命只有10^{-23} s左右，QGP火球的大小只有10 fm，实验上不能直接探测到QGP。高能核碰撞实验是典型的反问题，只能通过探测器捕获的末态粒子的四动量并结合高能核碰撞的理论模型仿真来反向推断初态核结构和核物质的性质信息。高能核碰撞理论相当于建立一个前向模型，给定不同的初始条件，核物质输运系数和状态方程，这个前向模型能够快速模拟一次核核碰撞，这个模型可以认为是核碰撞实验的数字孪生系统。理论模型详细考虑了高能核碰撞过程中的所有可能发生的物理过程，并用已知最好的理论来对其进行描述。比如，在美国布鲁克海文国家实验室，金原子核被加速到99.99%倍光速，在欧洲核子中心，铅原子核被加速到99.999 9%倍光速以上，这种情况下会发生极端的相对论效应，即尺缩与钟慢。在实验室参考系观察，原子核在束流方向的直径分别被压缩70和700多倍，原子核内部的时间演化也变慢70和700多倍，核子中的真空涨落没有足够时间回到真空，碰撞发生时不是几个价夸克之间的稀疏碰撞，而是大量海夸克与海胶子的密集碰撞。这对应着色玻璃凝聚以及胶子饱和的初态条件。当QGP达到局部热化，大量夸克和胶子之间的多体相互作用可以用相对论流体力学描述，此时QGP的温度高达几百兆电子伏特，真空中时刻存在大量的夸克反夸克对的激发和湮灭，夸克的数量、反夸克的数量、正电荷的数量、负电荷的数量不能单独守恒。此时只能使用基于能动量守恒、净重子数守恒、净电荷数守恒的相对论流体力学来模拟。模拟的过程需要知道QGP达到局部热化的时间，它的初态分布与涨落，QGP在不同温度下的剪切黏滞与体黏滞，QGP的状态方程（即压强作为温度和重子化学势的函数），QGP冷却转化为强子气体的温度等信息。当QGP转化为强子共振气体之后，还需要使用相对论的分子动力学模型继续模拟大量强子之间的碰撞、共振产生和衰变过程。可以看到，高能核碰撞仿真的标准模型非常复杂，涉及色玻璃凝聚、相对论流体以及强子输运几个过程中的众多参数，使用不同的模型参数，核碰撞仿真可以产生不同的末态强子分布。但如何设置

模型参数组合才能让理论模型描述实验观测数据是很困难的问题。甚至这有可能是一个病态的反问题，即很多种不同的参数组合都可以描述实验数据，此时通过实验数据无法唯一地确定高能核碰撞的初态性质以及QGP的输运系数、状态方程等信息。如图1[4]所示，理论模型的输出依赖于除了状态方程之外的众多参数，而每个参数可以影响多个实验的可观测量。这种多个模型参数与多种实验可观测量之间的复杂耦合为确定核物质状态方程或其他模型参数带来极大困难。这种情况的出现一部分原因是高能核碰撞末态数据经过了压缩和降维，实验提供的可观测量并没有包含所有的核碰撞末态信息；另一个原因是末态粒子谱与核碰撞过程中的多个参数耦合在一起，同一个物理参数可以影响几个可观测量，同一个可观测量对多个物理参数敏感。这种模型参数与末态可观测量之间的复杂耦合使得实验与理论的对比异常艰难。

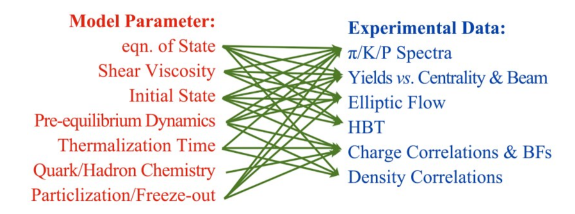


图1 通过高能核碰撞实验与理论模型对比提取核物质状态方程所遭遇的困难[4]

Fig.1 Demonstration of difficulty encountered when extracting the nuclear equation of state (EoS) through model-data comparison in heavy-ion collisions[4]

高能核碰撞理论与实验结果对比的另一个困难是非平衡态演化过程中的熵增。熵是系统不确定度的期望值。在剪切黏滞系数不为零的相对论流体力学演化以及强子输运过程中存在熵增，或者说信息量的丢失。熵增使得系统不可逆，即因为信息量的丢失而无法从末态反演到初态，也无法通过核碰撞末态所有粒子的四动量反推出初态核结构以及演化过程中QGP的全部信息。熵增使得高能核碰撞理论存在很大的不确定性，即很难从理论上确定QGP的某个重要物理参数的信息是否在剧烈的非平衡态演化过程中幸存了下来，并编码在末态强子的动量分布以及两粒子或多粒子关联中。熵增也为从末态粒子分布解码初态核结构和QGP的性质带来额外的困难，即理论反演的存在性困难。

机器学习中的贝叶斯推断技术以及深度学习技术为求解高能核碰撞物理中的反问题提供了巨大帮助[5]。贝叶斯推断技术通过对给定实验数据时模型参数的后验分布采样，可以获得全局拟合情况下的理论模型参数的分布，从而得到最优参数组合以及不同参数的不确定度。深度学习技术推动了当前的人工智能热潮，它非常擅长在不同的数据源之间建立映射。高能核碰撞理论模型可以获得不同物理参数设置下的末态粒子分布，即在前向过程中建立参数数据源与末态粒子数据源的映射。而深度学习中的监督学习技术可以通过训练多隐藏层人工神经网络，在末态粒子数据源与参数数据源之间建立逆向映射。建立逆向映射的过程是从零开始逐渐尝试的过程，如果最终尝试成功，说明物理参数的信息没有因为熵增而完全消失，深度学习技术可以把物理参数的信息从末态粒子分布中解码出来；如果尝试失败，说明物理参数的信息要么因为熵增而消失，要么深度神经网络的表示能力还不足以建立两者之间的非线性映射。深度学习从一定程度上解决了理论反演的存在性难题。深度学习除了解决理论反演的存在性难题，它还同时给出了理论反演的答案。使用机器学习中的可解释性研究，可以找到深度神经网络输入数据中对最终决策起到决定性作用的区域。这些区域对应着输入图片中的多部位关联，或者输入粒子云中的多粒子关联。在AI for Science研究领域，可解释性研究为物理学提供了很多启发。比如，根据输入图片或点云中起决定性作用的区域，可以获知最重要的动量区间。

高能核碰撞的一个重要研究目标是确定不同温度(T)和重子化学势(μ_B)条件下QGP向普通核物质以及强子共振气体转变的量子色动力学(Quantum Chromodynamics，QCD)相变类型。温度是核物质中组分粒子的平均热运动动能，重子化学势反映了核物质中的正反物质比例。比如，当原子核以接近光速运动并对撞时，入射核与靶核如透明般互穿，只在穿过的中心快度区域激发出大量的胶子、夸克与反夸克。这个区域的正物质数量(夸克)与反物质(反夸克)数量基本相等，对应零重子化学势。另外，宇宙早期的温度非常高，从真空中激发出大量的正反物质，正反物质数量基本相等，也对应着低重子化学势。QCD相图中的高温低重子化学势区域对应欧洲核子中心的超相对论重离子碰撞实验以及宇宙早期，Lattice QCD预测这个区域QGP到强子共振气体的相变为平滑过渡，即此区域热力学量的无穷阶导数连续。在中低能核碰撞中，入射核与靶核中的核子(正物质)会沉积在碰撞重叠区域，数量远多于从真空中激发出的等量正反物质，因此，对应QCD相图中的中间温度和高重子化学势区域。很多理论模型预测这个区域的核物质相变种类为一阶相变，一阶相变的端点被称作临界点。

本文集中介绍在重离子碰撞领域，机器学习如

何帮助确定核物质的状态方程，关于机器学习在物理以及核物理中更加广泛的应用，可以参考文献[6–14]。本文内容包括：1）使用Lattice QCD研究核物质状态方程以及相变种类遭遇的困难；2）使用贝叶斯分析方法研究核物质相变和状态方程；3）深度神经网络方法介绍，以及区分核物质状态方程常用的卷积网络，点云网络和多事件平均技术；4）使用动态棱卷积图神经网络寻找临界自相似；5）使用Lattice QCD状态方程学习准粒子质量；6）主动学习排除不稳定的参数空间；7）核液气相变；8）低能核碰撞，确定对称能；9）马赫锥与核物质状态方程的关系以及使用机器学习寻找马赫锥；10）辅助Lattice QCD采样，帮助解决Lattice QCD的符号问题。

1 Lattice QCD的符号问题

在离散的四维时空写下QCD的配分函数：

$$Z(\beta, N_\sigma, N_\tau) = \int \mathcal{D}U\mathcal{D}\Psi\mathcal{D}\bar{\Psi}\mathrm{e}^{-S_E(U,\Psi,\bar{\Psi})} \tag{1}$$

其中：$\beta = 1/T = N_\tau a$表示温度的倒数；N_τ是离散化的时间格点数；a是离散化的时空格子大小；N_σ表示每个空间维度的离散化格子数。$U_{x,\mu}$表示SU(3)空间的规范场，用x,μ标注格子之间的规范链，Ψ和$\bar{\Psi}$表示夸克与反夸克场。其中$S_E = S_G + S_F$，S_G和S_F分别表示胶子和夸克的作用量。此配分函数相当于对夸克场和胶子场做路径积分，费米子自由度可以解析积掉。不同的夸克场和胶子场构型（Field configurations）对应不同的作用量，以及不同构型出现的概率。配分函数是所有可能的场构型的非归一化概率之和。使用马尔科夫链蒙特卡罗技术可以从上述非归一化的概率分布采样，获得大量的夸克场和胶子场构型。使用这些采样出的夸克场和胶子场，可以计算零重子化学势和有限温度下核物质的状态方程，即压强，能量密度，熵密度以及迹反常（Trace anomaly）与温度之间的关系。比如，有了这些构型，可以获得迹反常：

$$\Delta_{\mu\nu} = \frac{\epsilon - 3P}{T^4} = T\frac{\mathrm{d}}{\mathrm{d}T}\left(\frac{P}{T^4}\right) = -\frac{V}{T^3}\frac{\mathrm{d}\ln Z}{\mathrm{d}\ln a} \tag{2}$$

使用式(2)还可以计算出压强与温度的关系以及能量密度与温度的关系，从而得到压强与能量密度之间的关系$P = P(\epsilon)$以及热密核物质的声速：

$$c_s^2 = \frac{\mathrm{d}P}{\mathrm{d}\epsilon} \tag{3}$$

但是，在有限重子化学势因为电荷共轭对称性的破坏，夸克场高斯积分中的虚数部分不能与反夸克场对应的积分相消，导致被积函数出现严重的振荡行为，形如$\int \mathrm{e}^{-x^2}\cos(\mu x)\mathrm{d}x$，其被积函数如图2所示，假设$\mu$代表重子化学势，纵轴表示被积函数。

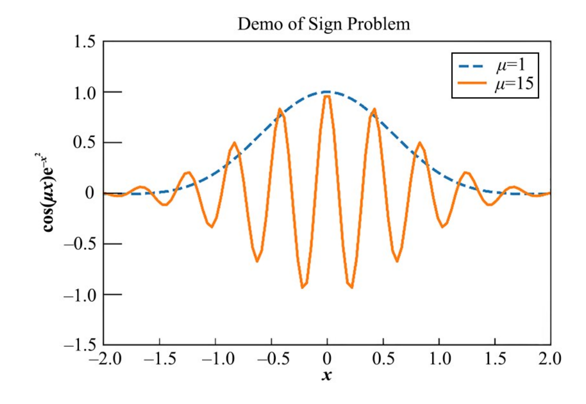


图2 格点QCD符号问题示意图
Fig.2 Demonstration of the sign problem in lattice QCD

在零重子化学势附近，可以对状态方程进行泰勒展开，获得低重子化学势区域的状态方程[15–16]。泰勒展开在化学势较大的区域自动失效，直接使用Lattice QCD又会遇到上面展示的符号问题，因此，目前基于非微扰QCD的第一性原理计算技术Lattice QCD只能探索核物质相图中的很小一个区域，无法给出临界点的存在性与具体位置信息。

2 使用贝叶斯分析从实验与理论的对比中确定核物质状态方程

根据简单的联合概率密度分布公式$P(x,y) = P(x|y)P(y) = P(y|x)P(x)$可以得到贝叶斯公式：

$$P(x|y) = \frac{P(y|x)P(x)}{P(y)} \tag{4}$$

在高能核碰撞物理中，x表示理论模型参数，y表示实验数据。贝叶斯分析的目标是在获得新的实验数据y的情况下，计算理论模型参数x的后验分布$P(x|y)$。此后验分布可以简单看作理论模型参数满足的概率密度分布函数，它有均值、有方差、有峰值。确定峰值所对应的x称作最大后验估计，方差表示理论模型参数的不确定度。因此，贝叶斯分析不仅可以帮助物理学家自动寻找最优的模型参数组合，而且可以将实验数据中的误差以及理论模型的不确定性反向传递到参数的后验分布中。贝叶斯公式中的$P(y|x)$被称作似然函数，它量化了参数为x的理论模型输出与实验数据y之间的似然程度。$P(x)$被称作模型参数的先验，即新的实验数据出现之前大家对模型参数可能取值的猜测。从公式可以看出，先验和似然都会影响模型参数的后验分布。分母上的归一化常数$P(y) = \int P(y|x)P(x)\mathrm{d}x$被称作证据，计算需要对整个参数空间，即$x$的所有可能取值

进行积分，因此计算量巨大。幸运的是，马尔科夫链蒙特卡罗技术可以对如下形式的非归一化的概率密度采样：

$$P(x|y) \propto P(y|x)P(x) \tag{5}$$

因此，贝叶斯分析技术其实是对理论模型参数x的非归一化后验概率密度$P(x|y)$分布采样的过程。使用马尔科夫链蒙特卡罗技术，可以从参数空间中任意一点x_0出发，根据细致平衡条件做随机行走：

$$x_{i+1} = x_i + \varepsilon \tag{6}$$

式中：ε是随机行走的步长，可以从均匀分布或高斯分布采样。行走的脚印，或者说参数空间中的历史轨迹形成一个集合$\{x_0, x_1, \cdots, x_n\}$，这个集合即是对后验概率密度分布的采样。对这些采样做密度估计，就可以得到后验分布的峰值位置以及方差。

格点QCD预言相对论重离子对撞机(Relativistic Heavy Ion Collider，RHIC)和大型强子对撞机(Large Hadron Collider，LHC)能量下的核核碰撞产生高温低重子化学势的夸克胶子等离子体，冷却强子化经历了一个平滑过渡(Crossover)。一个很重要的问题是：是否能在实验上验证QCD相图上的这个区域经历了平滑过渡？是否能够通过RHIC和LHC的大量实验数据确定这个区域核物质状态方程的形式，即压强作为能量密度的函数。为了解决这个问题，首先需要在高能核核碰撞的理论模型中确立核物质状态方程对末态可观测量的正向影响，这一步可以通过相对论流体力学实现[17-18]。零重子化学势下，相对论流体力学求解如下能动量守恒方程：

$$\nabla_\mu T^{\mu\nu} = 0 \tag{7}$$

其中：能量动量张量$T^{\mu\nu} = (\epsilon + P)u^\mu u^\nu - P g^{\mu\nu} + \pi^{\mu\nu}$，$\epsilon$是能量密度，$P$是压强，$u^\mu$是流体的四速度，$\pi^{\mu\nu}$是剪切黏滞张量。相对论流体力学公式描述了夸克胶子等离子体以及高温高密的强子物质的时空膨胀演化。膨胀的驱动力主要是压强梯度和剪切黏滞。改变状态方程，即压强与能量密度的函数关系，会改变热密核物质的膨胀加速度。在相对论流体力学的框架下，Pratt等[19]构造了一组参数化的状态方程，通过理论模型与实验数据的对比，计算出核物质状态方程参数的后验分布。参数化的形式如下：

$$c_s^2(\epsilon) = c_s^2(\epsilon_h) + \left(\frac{1}{3} - c_s^2(\epsilon_h)\right)\frac{X_0 x + x^2}{X_0 x + x^2 + X'^2} \tag{8}$$

图3是使用贝叶斯分析提取出的高温度零重子化学势区域的核物质状态方程。不难看出，未做约束之前，状态方程参数的先验分布接近均匀分布，参数取值有很大的不确定度，导致状态方程以及声速的平方有很多可能的取值。通过实验与流体力学输出结果的对比，可以计算出似然函数。通过贝叶斯分析，计算出参数满足的非归一化后验分布并采样，最终获得符合高能核碰撞实验数据的核物质状态方程。可以看到，经过实验数据的约束，状态方程的可能取值范围缩小到一个很小的区域，且与Lattice QCD预测的高温低密核物质状态方程相匹配。在使用贝叶斯分析拟合其他模型参数的过程中，Lattice QCD的状态方程一般直接作为输入[20-22]。最近，文献[23]使用贝叶斯方法，通过对重离子碰撞产生的末态质子的平均横动能量以及椭圆流的分析，从而有效地约束了高密区的QCD状态方程。

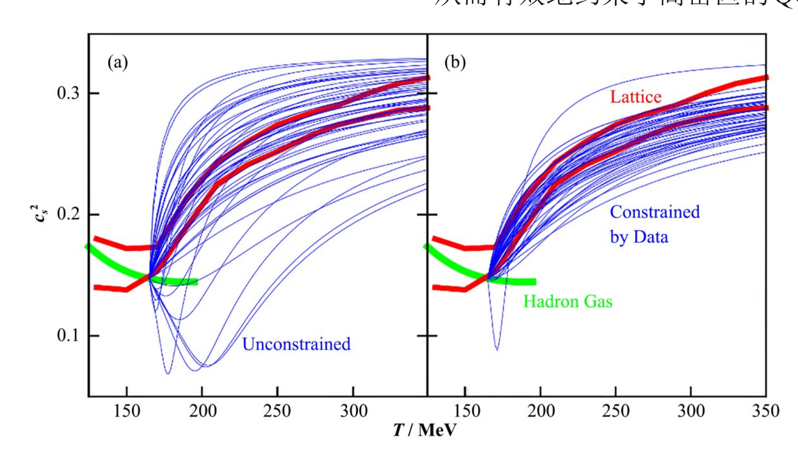


图3 使用贝叶斯分析提取出的高温零重子化学势区域的核物质状态方程[19]

Fig.3 Extraction of nuclear EoS at a high temperature and zero baryon chemical potential region using Bayesian analysis[19]

贝叶斯分析技术能够充分利用物理先验以及实验提供的低维数据，并把实验误差甚至是理论模型误差通过似然函数向后传递到参数的后验分布。这种技术的一个缺点是最终结果依赖于人工设定的参数化函数形式，可能会引入人为偏见。另一个缺点是高维实验数据向低维可观测量压缩投影的过程中损失了大量信息，贝叶斯分析技术不能帮助从原始数据中发现新的可观测量。

3 深度神经网络确定核物质状态方程

深度学习（Deep Learning）一般被理解为多隐藏层人工神经网络。深度学习领域的三驾马车之一LeCun将深度学习定义为构造层级化的泛函模块，并通过梯度下降对其优化的方法。根据这个定义，深度神经网络提供了将输入x非线性的映射到输出y的泛函模块，写成函数形式为$f(x,\theta)$，在监督学习中，如果知道标准答案y，则优化目标是神经网络输出与标准答案之间的距离，定义为损失函数：

$$L(\theta)=\mathrm{Distance}[y,f(x,\theta)]$$

损失函数L是函数f的函数，返回值为标量，所以称为泛函。优化的目标是通过调整神经网络参数θ的方式来逐渐减小神经网络输出与标准答案之间的距离。

接下来介绍神经网络的参数θ及其优化方式。考虑简单的线性变换，$z=xW+b$，其中：输入x是拥有m个元素的行向量，W是m行n列的矩阵，b是1行n列的矩阵，这样的矩阵操作返回一个新的n个元素的向量z。W和b中矩阵元的整体被记作θ。这个矩阵操作相当于把m维空间的向量通过旋转、缩放、平移以及升降维度的方式变换到n维空间。这种矩阵操作可以用如下的单层人工神经网络表示，第i个输入神经元与第j个输出神经元之间的连接权重的值等于矩阵元W_{ij}，每个输出神经元还要另外加上偏置b_j。此时，人工神经网络拥有简单的线性表示能力，可以用于线性拟合任务。单纯增加线性神经网络的层数不能增强其表示能力，原因是多个矩阵的乘积等于一个矩阵。将z的每一个元素都做一个非线性变换$h=\sigma(z)$，则会为人工神经网络带来更强大的表示能力。原则上只要在输入和输出层之间添加一个隐藏层，在足够多神经元的情况下，神经网络就会拥有逼近任何连续函数的能力。图4为一个单层人工神经网络示意图。

神经网络的参数一般按特定的方式初始化为随机数，以避免多个矩阵连乘导致的梯度消失或梯度爆炸。初始时神经网络将输入x随机地映射到输出$f(x,\theta)$，对于给定的输入，神经网络做出随机预测，预测结果与标准答案存在较大的偏差。在深度学习中，一般使用随机梯度下降算法（Stochastic Gradient Descent，SGD）对神经网络中的每一个参数进行更新。最简单的SGD版本中，参数的更新方法为：

$$\theta_{n+1}=\theta_n-\varepsilon\frac{1}{m}\sum_i^m\frac{\partial L_i}{\partial\theta_n} \tag{9}$$

式中：ε是一个很小的正数，表示学习率；m表示一个小批次的样本数量；L_i表示第i个样本的损失函数。

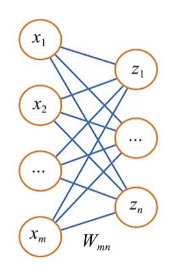
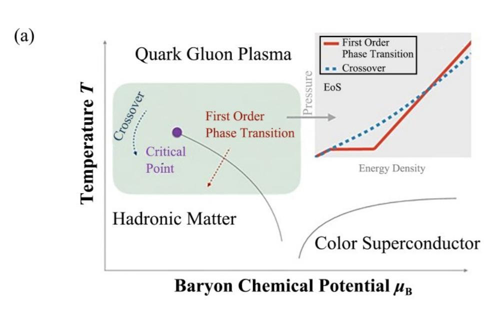
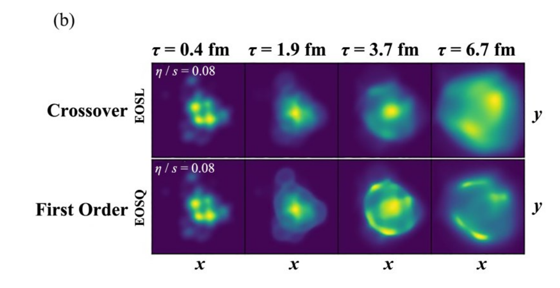


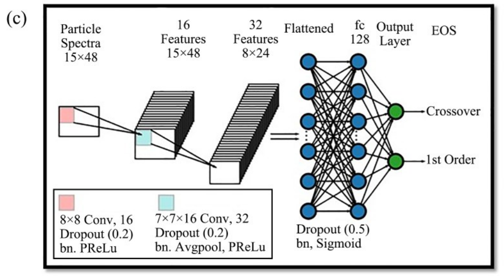
图4 单层人工神经网络对数据的变换等价于线性矩阵操作$z=xW+b$，再使用非线性激活函数将z的每个元素进行变换
Fig.4 A single layer artificial neural network transforms the input data by a linear matrix operation $z=xW+b$, in combination with a non-linear operation on each element of the output z using a non-linear activation function

使用损失函数的负梯度来更新每个参数的值，可以让损失函数的值逐渐减小。这一点可以通过简单的例子说明，如果$L_i=\theta_n^2$，则极小值的位置为$\theta_n=0$处，任意处负梯度为$-2\theta_n$，如果初始的$\theta_n>0$，使用$-2\varepsilon\theta_n$更新参数会让其减小，相反，如果初始化时$\theta_n<0$，则负梯度$-2\varepsilon\theta_n>0$，也会朝着损失函数减小的方向更新。中等大小的人工神经网络拥有上千万个可训练参数，神经网络学习的过程需要计算损失函数对每个参数的负梯度。数值的有限差分算法会带来不可控的截断误差与舍入误差，解析微分会遭遇表达式膨胀问题，计算效率低下。深度学习中的负梯度一般使用自动微分[12]（Auto Differentiation）计算，它是拥有解析精度的高效数值微分算法，被广泛用于深度神经网络的优化以及求解常微分方程和偏微分方程的Physics Informed Neural Network中。

文献[24]开创了使用深度学习这种强大的模式识别工具判断核物质状态方程以及相变种类的先河。如图5(a)所示，QCD相图上平滑过渡区（Crossover）对应的状态方程用蓝色虚线表示，即压强作为能量密度的函数在高能量密度区（夸克胶子等离子体相）与低能量密度区（强子共振气体相）之间是光滑的插值函数，标记为EOSL。一阶相变区（First order phase transition）对应的状态方程用红色实线表示，标记为EOSQ，压强P作为能量密度ϵ的函数在相变区域存在平台结构，在此平台上压强梯度$\mathrm{d}P/\mathrm{d}\epsilon=0$，QGP的膨胀加速度为0，出现集体流骤减壳层。反映在图5(b)中，使用EOSL，热密核物质时空演化过程中能量密度分布更加平滑均匀，使用EOSQ，因为一阶相变区的影响，能量密度分布存在更多的山岭状结构。QGP集体流动过程中的区别会反映在最终末态强子的动量分布上。这个研究展示了使用传统的物理可观测量，无法根据每个事件的末态粒子谱区分QGP演化时遵循的状态方程。







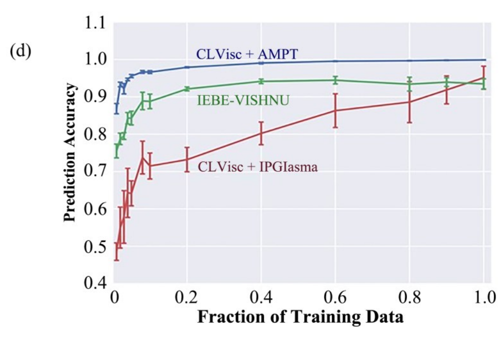


图5 使用深度卷积神经网络从末态粒子分布提取核物质状态方程以及相变种类信息[24]

Fig.5 Identification of the nuclear EoS and types of phase transition from the final state particle spectra using deep convolutional neural network[24]

既然深度学习可以从大量的猫狗图片中提取最有用的特征或特征组合，实现对不在训练集中的猫狗分类，那它是否可以被用来从相对论重离子碰撞逐个事件的末态粒子分布中提取特征组合，从而判断夸克胶子等离子体到强子共振气体之间的相变种类？文献[24]使用监督学习的思路，通过相对论流体力学模拟产生大量的高能核碰撞事件，使用两种状态方程计算出的末态强子谱作为神经网络的输入，将状态方程种类作为神经网络的输出，训练卷积神经网络在输入与输出之间建立映射。在这个研究中，输入数据为15×48个像素的二维粒子谱$\rho(p_T,\phi)$，表示15个不同的横动量以及48个不同方位角处的粒子数密度。输出层有两个神经元，取值都在(0,1)之间，和为1，表示平滑过渡和一阶相变的归一化概率。这一步使用softmax激活函数实现：

$$q_i=\frac{e^{z_i}}{\sum_j e^{z_j}} \tag{10}$$

分类任务一般使用交叉熵损失函数，来量化两个概率密度分布之间的差别，

$$L=-\sum_i p_i \ln q_i \tag{11}$$

式中：p是标准答案对应的概率密度分布；q是神经网络预测出的概率密度分布；q_i表示神经网络预测输入数据的标签是第i类的概率。如果真实标签是平滑过渡，则：

$$p=\begin{pmatrix}1\\0\end{pmatrix}, q=\begin{pmatrix}q_{\rm EOSL}\\q_{\rm EOSQ}\end{pmatrix} \tag{12}$$

如果真实标签是一阶相变，则：

$$p=\begin{pmatrix}0\\1\end{pmatrix}, q=\begin{pmatrix}q_{\rm EOSL}\\q_{\rm EOSQ}\end{pmatrix} \tag{13}$$

使用AMPT（A Multiphase Transport）模型产生逐个事例相对论流体力学模拟的初始条件，使用两种状态方程，几个不同的剪切黏滞系数以及冷却强子化温度，进行相对论流体力学模拟，来使数据尽可能地多样化。如图5(d)所示，经过训练的人工神经网络在三组不同的测试数据集上都能达到93%以上的测试精度。其中IEBE-VISHNU[17]使用不同的初始条件和流体力学代码，CLVisc+IPGlasma[25]使用与训练数据相同的流体力学代码和不同的初始条件，体现出模型良好的泛化能力。模型数据与真实实验数据之间可能存在偏差，为了减小此偏差，后续的研究一方面在模型中加入强子输运，考虑QGP转化为强子后的级联散射、共振产生和衰变过程[26]，另一方面考虑到探测器的有限接收度，将仿真数据进行进一步的探测器模拟修正[27]。这些研究发现，如果末态粒子存在比较大的涨落，多事件平均的技术可以显著增加分类准确度。

高能核物理以及粒子物理的末态粒子数据结构为天然的动量空间的点云。如果使用卷积神经网络，需要对粒子进行密度估计，将粒子的动量投影到(p_x,p_y)或(p_T,p_φ)等二维离散空间，产生图片。这个过程会产生精度丢失，纵向信息损失，以及严格守恒律的破坏。如果保留纵向动量，则三维格子中的粒子信息非常稀疏。点云神经网络直接使用每个粒

子的四动量以及其他量子数作为输入，精确的保留每个粒子的完整信息，更加适用于高能核物理以及粒子物理。

文献[28]使用点云神经网络区分一阶相变中的Spinodal类型和Maxwell类型状态方程。其中Spinodal状态方程存在一个声速平方$c_s^2 = \mathrm{d}P/\mathrm{d}\epsilon < 0$的不稳定区域，这个区域会导致QGP形成多个分离的区块，每个区块独立强子化。Maxwell构造对应之前介绍的EOSQ状态方程，存在压强梯度为0的区域。如图6所示，这个研究将一个全连接网络应用于每个粒子，将其四动量以及质量转化为128维高维特征，此网络被所有粒子共享，又称一维卷积。全局最大池化Global Max Pooling被用来寻找粒子云在128维特征空间中的边界，这些边界被输入到另一个全连接神经网络，最后输出一个神经元，使用Sigmoid激活函数，使输出满足$0 < y < 1$。其中$y > 0.5$表示预测结果为Spinodal，$y < 0.5$表示预测结果为Maxwell。此外，文献[29]还使用点云神经网络识别QCD状态方程中的一阶相变模式与平滑过渡模式，其分类的准确度高达99.8%。

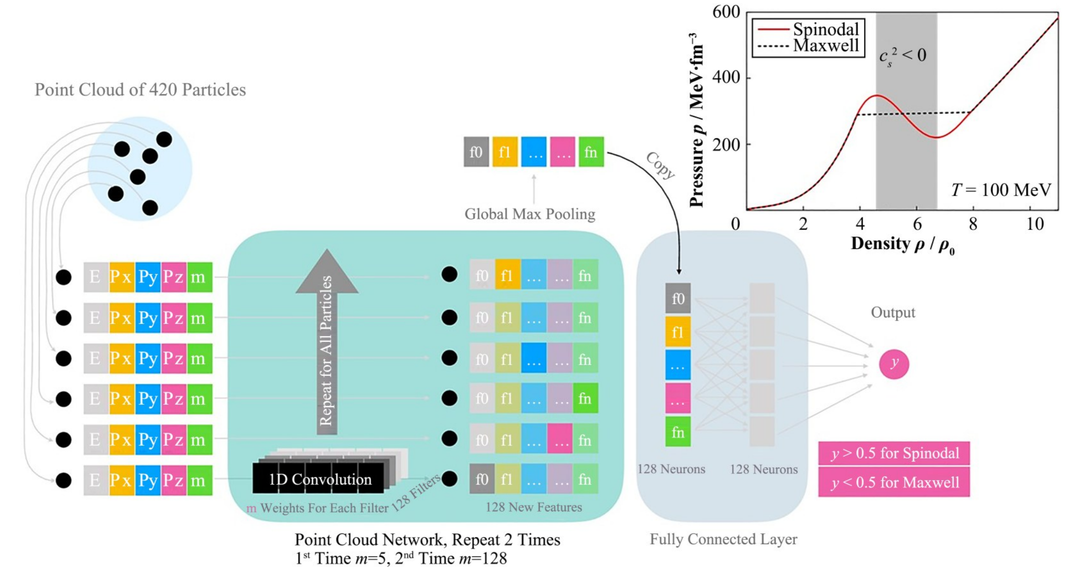


图6 使用点云神经网络从动量空间的末态粒子云提取核物质状态方程信息[28]

Fig.6 Identification of the nuclear EoS from the final state particle cloud in momentum space using point cloud neural network[28]

4 使用深度学习帮助寻找相变临界点

除了利用一阶相变区集体膨胀加速度变小的特点，还可以使用临界点附近涨落的增强现象研究核物质相变。在临界点附近，关联长度发散，出现临界乳光和标度不变。后者在物理上体现为不同分辨率下观察系统出现类似分形的自相似现象。文献[30]使用卷积神经网络从σ标量场的二维时空构型中判断是否存在QCD相变信息。文献[31−35]建议临界点附近的粒子可能满足动量空间的自相似，使用间歇性分析可以探测重离子碰撞中的自相似性和临界波动。尽管最近欧洲核子研究中心SPS的NA61/SHINE实验[36−38]在150 AGeV的Ar+Sc碰撞中观察到了一个不可忽略的间歇现象，但背景涨落很大，幂律关系无法完全确定。另外，相近碰撞能量下的C+C、Pb+Pb和Be+Be碰撞并没有观察到间歇性信号。临界蒙特卡罗模拟表明，在Be+Be碰撞中，临界质子的最大可能比例小于0.3%，这表明如果CMC（Critical Monte-Carlo）无粒子的分数与不相关的背景相比很小，那么传统的间歇性分析可能无法寻找到自相似性的微弱信号。使用间歇性分析寻找临界信号的另一个困难是探测器的动量分辨率较低，如果按照探测器的分辨率极限给蒙特卡洛模型产生的每个末态粒子添加5%的动量误差，则间歇分析无法区分加入了5%或10%临界关联粒子的事件。

文献[39]使用点云神经网络和动态棱卷积神经网络来寻找重离子碰撞中每个事件的微弱自相似信号。JAM（Jet AA Microscopic Transportation Model）蒙特卡罗模型被用来产生Au+Au$\sqrt{s_{NN}}$ = 27 GeV中心碰撞的22万个事件，这些事件的末态粒子不存在动量空间的自相似，标记为JAM。CMC模型被用来产生存在动量空间自相似的关联粒子，每对粒子在动量空间的间隙Δp满足如下幂律分布：

$$f(\Delta p) = A\Delta p^{-(1+\nu)} \tag{14}$$

其中：$\nu = \frac{1}{6}$是三维伊辛模型普适类对应的一个临界指数，A是与ν有关的一个归一化常数。这个概率密

度分布的累积概率函数可以解析求出，且有反函数，可以使用反函数法抽样Δp。从JAM粒子中任选一个动量，根据抽样得到的Δp，获得一系列有关联的CMC粒子。文献[39]将一部分JAM事件中5%或10%的粒子替换成存在临界自相似的CMC粒子，并将这类事件标记为CMC。点云神经网络和动态棱卷积神经网络被训练来区分两种事件，检验深度学习是否能从大量背景粒子中挑选出少数存在长程关联的粒子。

图7是动态棱卷积神经网络的架构。神经网络的输入是动量空间的粒子云。网络输出分为两部分：第一部分处理分类任务(Classification)，输出只有一个神经元，其取值可以区分一个事件中是否存在临界信号；另一部分处理标注任务(Tagging)，输出神经元的个数等于输入粒子的个数，其取值表示对应的输入粒子是信号粒子还是噪声粒子。输入到输出之间的变换使用动态棱卷积网络，这个网络首先寻找动量空间每个粒子的k个最近邻粒子，然后使用一个全连接网络计算每个粒子跟它一个邻居的edge信息：

$$e_{ij} = \mathrm{net}(p_i, p_j) \tag{15}$$

这个全连接网络被所有的粒子对共享，因此称为棱卷积。每个粒子的特征和k个edge信息拼接形成一个向量$[p_i, e_{i1}, e_{i2}, \cdots, e_{ik}]$，一起输送给点云网络，作进一步的处理。由此可知，相比点云网络，动态棱卷积网络使用了更多局部结构信息。需要注意的是，棱卷积可以重复多次。第一次是在动量空间寻找每个粒子的k个最近邻粒子，从第二次开始就是在特征空间寻找每个粒子的k近邻粒子，这也是动态棱卷积名称的来源。特征空间距离相近的粒子在动量空间可能相隔甚远，因此，这种神经网络架构擅长从大量粒子中寻找存在长程关联或者多粒子关联的信号。经过训练，棱卷积神经网络寻找临界信号的能力相比点云神经网络提高10%左右，如表1所示。

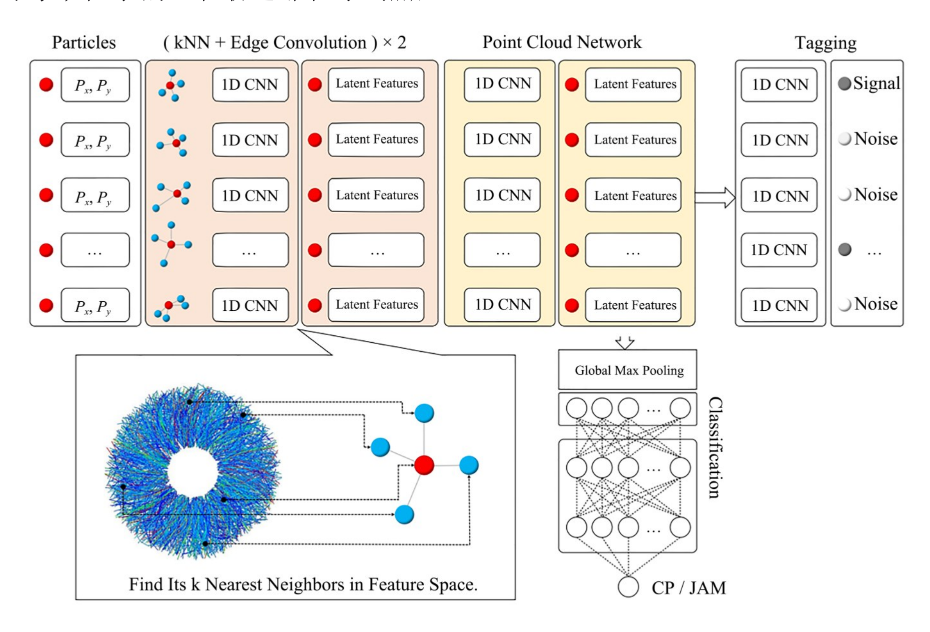


图7 使用动态棱卷积网络寻找动量空间的临界自相似信号，并标注存在临界关联的粒子[39]

Fig.7 Search for self-similarity in momentum space using dynamical edge convolution network and identification of correlated particles[39]

5 深度学习准粒子质量，重建Lattice QCD状态方程

零重子化学势下核物质状态方程可以通过Lattice QCD进行精确的计算，然而，计算所需的资源以及耗时非常昂贵。文献[40]中对比了Lattice QCD与简单的强子共振气体模型所计算的状态方程，结果表明无相互作用的强子共振气体模型能够很好地描述$T_c < 150$ MeV区域内的QCD状态方程，但是在高温区域却存在着较大的偏差。这是由于在高温区域，夸克和胶子不仅会从强子中解禁闭出来，还会从真空中激发出许多部分子。此时所产生的热核物质自由度主要为部分子，在相变区域附近存在很强的相互作用。文献[41]假设热密的QGP由非相互作用的准粒子组成，其质量取决于介质的温度。在物理学上，这些准粒子代表介质中的夸克和胶子，也被称作dressed粒子，相当于把相互作用势能吸收到每个准粒子的质量中。使用这些无相互作用但质

表1　棱卷积神经网络与点云神经网络寻找临界信号的性能对比

Table 1　Comparison between the performances of dynamical edge convolution network and point cloud network when searching for critical phenomenon

关联粒子比例 Associated particle ratio	动态棱卷积网络分类准确度 Classification accuracy of dynamic edge convolutional network	点云神经网络分类准确度 Classification accuracy of point cloud network
5%	92.8%	83.4%
10%	97.7%	84.8%

量依赖于温度的准粒子，作简单的统计力学计算，可以复现Lattice QCD的状态方程。

传统上，人们可以为温度依赖的准粒子质量提供一个参数化的函数$M(\theta,T)$，其中θ是由Lattice QCD数据所确定的参数[41]。然而，这种参数化的形式会引入一些人为先验的诱导性偏差。另外，这种方法也很难从QCD的状态方程中确定两个或者更多未知的准粒子质量函数。为了在不引入先验偏差的情况下解决这个变分问题，文献[42]采用深度神经网络（Deep Neural Networks，DNN）来表示三个未知的准夸克和胶子质量函数，以此作为夸克-胶子等离子体的局部温度的函数，并使用自动微分和梯度下降来解决这个问题。

图8是三个温度依赖的质量函数的神经网络模型。这三类准粒子质量模型分为上和下夸克$m_{u/d}(T,\theta_1)$、奇异夸克$m_s(T,\theta_2)$和胶子$m_g(T,\theta_3)$，其中θ_1、θ_2和θ_3是DNN的参数。每个质量模型代表一个残差神经网络（ResNet）模块。它由两个ResNet拼接而成，每一个ResNet有8个隐藏层，32个神经元和一个类swish的激活函数。在每个ResNet的末尾使用sigmoid激活函数。当网络输出了准粒子的质量之后，就可以计算他们的状态方程。对于一个无相互作用的准夸克和胶子气体，它的配分函数为：

$$\ln Z(T)=\ln Z_g+\sum_i \ln Z_{q_i}(T) \tag{16}$$

其中：$Z_g(T)$是来自胶子的贡献和$Z_i(T)$是来自味道为i的夸克。它们是通过对玻色-爱因斯坦分布和费米-狄拉克分布的动量积分计算出来的：

$$\ln Z_g(T)=-\frac{d_g V}{2\pi^2}\int_0^{\infty}p^2\,\mathrm{d}p\ln\left[1-\exp\left(-\frac{1}{T}\sqrt{p^2+m_g^2(T)}\right)\right] \tag{17}$$

$$\ln Z_{q_i}(T)=+\frac{d_{q_i} V}{2\pi^2}\int_0^{\infty}p^2\,\mathrm{d}p\ln\left[1+\exp\left(-\frac{1}{T}\sqrt{p^2+m_{q_i}^2(T)}\right)\right] \tag{18}$$

式中：V是系统的体积；p是准粒子的动量；T是温度；$m_g(T)$和$m_{q_i}(T)$是神经网络的输出。d_g和d_{q_i}是胶子和夸克的自由度。然后再根据统计公式：

$$P(T)=T\left(\frac{\partial\ln Z(T)}{\partial V}\right)_T \tag{19}$$

$$\varepsilon(T)=\frac{T^2}{V}\left(\frac{\partial\ln Z(T)}{\partial T}\right)_V \tag{20}$$

获得相应的压强P以及能量密度ε之后，计算它们的熵密度$s(T)=\dfrac{\varepsilon+P}{T}$和迹反常$\Delta(T)=\dfrac{\varepsilon-3P}{T^4}$。这项工作的主要目标是确定$m_{u/d}(T,\theta_1)$、$m_s(T,\theta_2)$和$m_g(T,\theta_3)$，方法是使网络预测与Lattice QCD计算的$s(T)$和$\Delta(T)$之间的均方误差最小。在神经网络的向前传播过程中，采用高斯-拉盖尔求积法对配分函数公式(17)和(18)进行数值积分。而式(19)和(20)中所示的配分函数的导数由auto-diff技术自动提供，原则上具有解析精度。

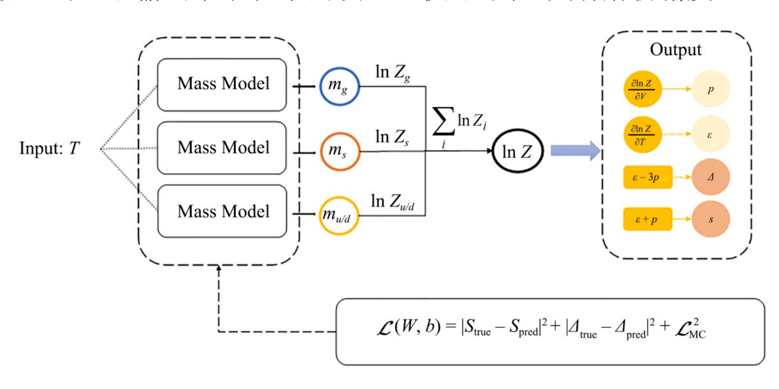


图8　使用深度神经网络构造3个温度依赖的质量函数，并且计算配分函数从而获得QCD的状态方程[42]

Fig.8　Framework of the deep neural network when constructing three temperature-dependent mass functions and calculating the QCD EoS using DNN learned masses[42]

训练目标是最小化损失函数，其中包含两个部分：一个是网络输出的熵密度和迹反常与Lattice QCD计算之间的RMSE（Root Mean Square Error）；另一个是约束来自于高温时的部分子有效质量。例如，硬热圈计算显示，在高温情况下，轻夸克（u、d）与胶子的有效质量比值为：

$$R_{g/q} = \frac{M_{g,T>2.5T_{\rm cut}}}{M_{q,T>2.5T_{\rm cut}}} = \sqrt{\frac{3}{2}\left(\frac{N_C}{3}+\frac{N_f}{6}\right)} \tag{21}$$

其中：高温截断为$T_{\rm cut} = 0.150$ GeV。那么$m_{u/d}(T,\theta_1)$和$m_g(T,\theta_3)$之间的约束可以写为：

$$\mathcal{L}_1 = \left| R_{g/q} - \frac{3}{2} \right| \tag{22}$$

对于奇异夸克，在高温下的约束条件被设定为：

$$\mathcal{L}_2 = \left| \frac{m_s - m_{u/d}}{\overline{m_s} - \overline{m_{u/d}}} - 1 \right| \tag{23}$$

式中：$\overline{m_s}$与$\overline{m_{u/d}}$为各自的流质量。那么，总的损失函数为：

$$\mathcal{L} = \left(s_{\rm true} - s_{\rm pred}\right)^2 + \left(\Delta_{\rm true} - \Delta_{\rm pred}\right)^2 + \mathcal{L}_{\rm MC} \tag{24}$$

经过随机梯度下降算法（Adam）的训练，深度神经网络学到的质量函数不仅再现了$s(T)$和$\Delta(T)$，还再现$\varepsilon(T)$和$P(T)$。利用HotQCD和WB的Lattice QCD组的数据，DNN分别学习了两组质量函数，都可以很好地描述格点数据。这种携带物理约束的多任务的深度神经网络DNN显示了其强大的表征能力，可用于各种变分问题。

6 使用主动学习（Active learning）排除状态方程中不稳定区域

主动学习（Active learning）是与监督学习、非监督学习、半监督学习、自监督学习以及强化学习并列的一种机器学习方法。在深度学习时代，监督学习需要海量的带标注数据，而主动学习可以在少样本上更加高效地学习。它能够主动挑选重要的样本进行人工标注，然后使用少量标注好的数据进行监督学习。主动学习的流程如图9所示，学习的过程中存在两组数据集，一组为少量的已标注数据集，记为Labelled Data，另一组为大量的无标注数据集，只有数据没有标签，记为Unlabelled Pool。学习开始时，使用Labelled Data训练人工神经网络，然后拿预训练好的人工神经网络来预测Unlabelled Pool中的样本。对于简单样本，神经网络非常自信，以极高的概率将其归为一类；对于复杂样本，神经网络不确信，比如两分类任务中，它预测一个样本属于第一类的概率为51%，属于第二类的概率为49%，此时这个样本被称作uncertain样本，从Unlabelled Pool中剔除，进行人工标注后放入Labelled Data。使用标注好的数据集在下一轮监督学习训练时，神经网络就能从困难样本中学到更多特征。挑选uncertain样本的时候，可以把预测结果的熵作为判据，即选择$s = -\sum_i p_i \log p_i$大的样本，也可以多训练几个人工神经网络，形成评审团，让它们对同一个样本分别作出判断，将评审团成员最不能达成一致的样本进行人工标注，还可以给数据一个假想的标签，计算$\partial L/\partial\theta$，根据梯度选择对神经网络参数影响大的样本进行标注。

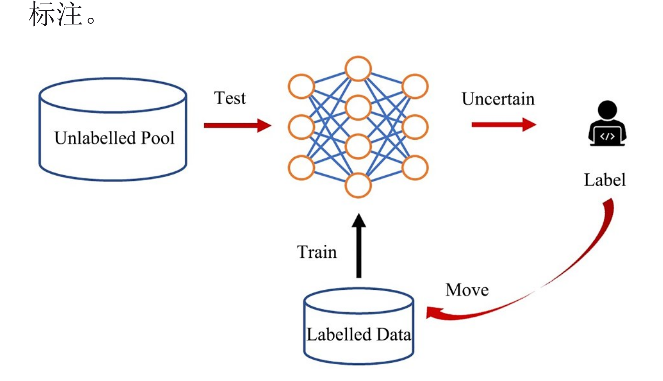


图9 主动学习（Active learning）的流程图[43]
Fig.9 Flow chart illustrating active learning[43]

文献[43]介绍了如何使用主动学习排除核物质状态方程中的热力学不稳定区域。束流能量扫描理论合作组（BEST）构造了一种存在相变临界点的状态方程[44]。构造的方法是将包含临界点的三维伊辛模型普适类对应的状态方程与零重子化学势附近做泰勒展开的格点QCD状态方程加权求和。这样构造的状态方程中临界区域在核物质相图中的大小和位置依赖4个参数，有些参数组合给出的状态方程违背热力学的稳定性与因果律。使用很小的参数空间，进行主动学习训练，人工神经网络可以在整个参数空间以96%的准确度判断出状态方程是否物理。主动学习技术在寻找合适的状态方程任务中，显著降低了机器学习模型对专家标注数据的需求，并增加了训练速度与准确度，可能会为人工智能与前沿科学的交叉带来更多突破。

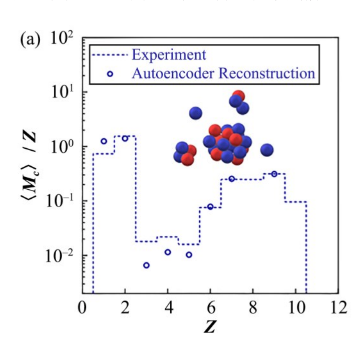
7 非监督学习核物质液气相变

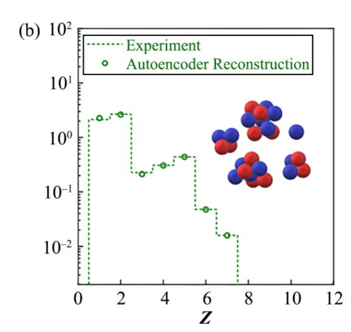
之前介绍的工作使用监督学习的方法，通过理论模型准备末态粒子谱或粒子云输入数据，将理论模型中使用的状态方程或经历的相变类型作为标签，训练机器学习模型在输入数据与标签之间建立映射。训练完成之后，机器学习模型可以在实验数据上进行预测。这种方法的局限是理论模型无法完

全描述实验数据，两者之间可能存在偏差。而直接使用实验数据作监督学习又会遇到缺少标注的困难。

非监督学习是一种机器学习模式，它可以自动地从无标注的数据中发现规律。线性与非线性降维Dimensionality reduction，聚类clustering，自编码机Auto Encoder，生成对抗网络GAN都属于非监督学习的范畴。非监督学习的一种思路是使用更少的信息重建原始数据。比如，非监督学习中的主成分分析（Principal Component Analysis，PCA），可以提取数据中方差最大的几个维度作为主成分，用很少的数据量表示原始数据中最重要的信息。PCA被用来分析末态粒子在动量空间的各向异性流，研究发现通过主成分分析提取的各向异性流与初始时刻能量密度分布的几何各向异性存在更佳的线性映射关系，显示相对论流体力学对不同模式的非线性耦合并没有人们预期的强[45]。自编码机一般使用两边大中间小的瓶颈结构，将输入映射到输出，损失函数定义为输出输入之间的距离，即重建误差。经过学习，自编码机瓶颈处的神经元存储着输入数据中最重要的信息。可以认为带瓶颈的自编码机起到非线性降维作用。

文献[46]使用非监督学习中的自编码机，从无标注的重离子碰撞实验数据中学习，进行核液气相变分类。如果碰撞能量很低，原子核被激发后处于单极震荡并偶发轻粒子，此时处于液相，碰撞能量增加后，原子核激发后进入spinodal区，因为核子核子之间的吸引相互作用形成中等质量的核碎片，此时处于液气相变区，继续增加碰撞能量，原子核会快速越过spinodal区，时间不足以形成中等质量核碎片，最终发射出大量轻原子核。如图10所示，不同碰撞能量以及物态对应的核碎片的电荷分布有明显区别。文献[46]将核核碰撞末态的带电粒子碎片的产额乘以其电荷数，即$ZM_c(Z)$作为神经网络的输入，使用带瓶颈的自编码机来重建输入分布。实验数据用图10中的虚线表示，自编码机的重建结果用空心圆点表示，两者较好符合。训练完成之后，瓶颈处的神经元取值出现明显的序参量特征，可被用来区分核液气相变种类。





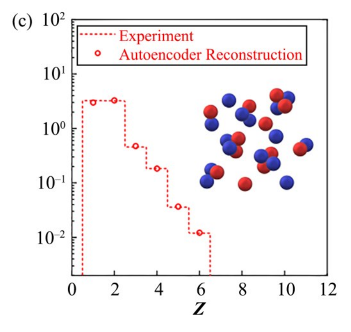


图10 使用自编码机从重离子碰撞实验数据中学习，进行核液气分类[46]
Fig.10 Classification of nuclear liquid gas phase transition using an auto encoder, which learns from the experimental data of heavy-ion collisions[46]

8 确定核对称能

在中低能重离子碰撞中，核物质状态方程作为能够表征核物质能量在不同的密度、温度以及同位旋不对称度的关系，一直以来都是理论与实验研究的热点之一文献[47−53]。对于理论而言，状态方程是非常重要的初态输入量，直接影响了重离子碰撞的动力学演化过程。对于同位旋非对称核物质，在不考虑温度时，其状态方程一般可描述为：

$$E(\rho,\delta) = E(\rho,0) + E_{\rm sym}(\rho)\delta^2 + O(\delta^4) \tag{25}$$

其中：$\delta = (\rho_n - \rho_p)/(\rho_n + \rho_p)$为同位旋不对称度；$E(\rho,0)$与同为旋无关为对称核物质能量；$E_{\rm sym}(\rho)$为对称能项。对称能作为构成状态方程非常重要的部分，如何通过末态观测量对他进行约束，从而获得核物质状态方程也引起了众多核物理工作者的关注[51−52]。对称能的通常表达式为：

$$E_{\rm sym}(\rho) = S_0 + L\left(\frac{\rho-\rho_0}{3\rho_0}\right) + \frac{K_{\rm sym}}{2}\left(\frac{\rho-\rho_0}{3\rho_0}\right)^2 + O(\rho^3) \tag{26}$$

其中：$L = 3\rho\left.\frac{\partial E_{\rm smy}(\rho)}{\partial\rho}\right|_{\rho=\rho_0}$ 表示饱和密度处的斜率参数；$K_{\rm sym} = 9\rho\left.\frac{\partial^2 E_{\rm smy}(\rho)}{\partial\rho^2}\right|_{\rho=\rho_0}$ 为曲率参数。

目前，理论上所给出的高密度区域的对称能仍存在很大的差异，但是对于末态粒子分布的影响不

十分明显。文献[54]采用了卷积神经网络的方法，通过对末态质子谱的学习，能够有效识别出初态的对称能的类型。其中，逐事件的二分类精度最高能达到60%左右，而100个事件叠加的二分类最高能达到98%左右，并且还采用100个事件叠加的方式进行了五分类的测试，其分类的精度在末态质子谱上能达到58%，在中子谱上能到达72%。该研究团队还进行了L的回归测试，在使用质子和中子的情况下，CNN(Convolutional Neural Networks)预测的L和真实的L之间的平均绝对误差分别为20.4 MeV和14.8 MeV。此外，文献[55]还使用了LightGBM算法。通过对所构造的30个末态特征量的逐事件进行学习，LightGBM能够将对称能斜率参数L的平均绝对误差控制在30 MeV左右，并且模型还具有很强的鲁棒性和泛化能力，如图11所示。

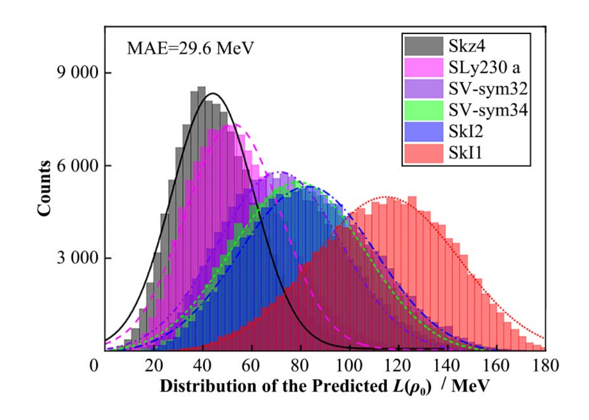


图11 LightGBM预测的斜率参数L的分布情况的分布[55]
Fig.11 Distribution of the slope L parameters predicted by LightGBM[55]

9 深度学习寻找QGP中的马赫锥

介质中高速运动的物体由于与介质的相互作用损失能量，比如子弹在水中飞行1 m左右即可失去杀伤力。如果物体的飞行速度大于介质中的声速，就会产生音爆，在飞行方向压缩出马赫锥，在后方留下耗散尾流。相对论重离子碰撞产生的高能部分子穿过QGP时会与热密介质中的热夸克和胶子发生强相互作用，通过弹性碰撞和胶子辐射损失能量。损失的能动量沉积在QGP中，带来的扰动在QGP中以声速传播。因为高能部分子的速度接近光速，远超QGP中的声速$c_s \approx 1/\sqrt{3}$，预期会在QGP中产生马赫锥和耗散尾流[56]。马赫锥的半锥角θ与声速c_s和部分子的飞行速度c之间存在简单三角函数关系$\sin\theta = c_s/c$。如果能在重离子碰撞实验中寻找到马赫锥的信号，测量到马赫锥的半锥角θ，则可以通过公式$c_s^2 = \mathrm{d}P/\mathrm{d}\varepsilon$直接获得热密核物质的状态方程信息。另外，马赫锥锥前的宽度与核物质的剪切黏滞也比较敏感[57]。但是，寻找QGP中的马赫锥信号异常艰难，因为马赫锥的形状会受到喷注的初始产生位置以及集体流的共同影响。

文献[58]使用机器学习辅助的喷注层析技术研究高能核碰撞中产生的马赫锥。如图12(b)所示，朝同一个方向出射的高能部分子喷注，如果在QGP中初始产生位置不同，则最终马赫锥的形状会同时受到局部温度和流速的影响。在高温区，介质中夸克和胶子的密度更高，喷注与介质的散射概率增大，马赫锥偏向高温区。但另一方面，集体流沿着温度梯度的方向，又会导致马赫锥向低温区偏转。两个效应存在竞争，文献[58]使用线性玻尔兹曼输运程序LBT(Linear Boltzmann Transport)[59]和相对论流体力学模拟程序CLVisc[18]来产生末态粒子分布与初态喷注产生位置训练数据集。图12(a)展示了用于喷注初始产生位置预测的点云神经网络架构。神经网络的输入包含两部分，第一部分是动量空间的末态粒子云，通过一维卷积将每个粒子的特征变换到高维特征空间，并使用全局最大池化获得粒子与粒子之间的相对关系，第二部分是喷注的全局信息，比如在光子标定的喷注中，光子的横动量，快度和方位角以及使用fastjet[60]获取的喷注的横动量，快度和方位角被送入一个全连接网络(Multilayer Perceptron，MLP)进行特征变换。两部分输入被拼接之后，再送入最后的回归网络，输出两个数字(x，y)，用来表示高能部分子喷注的初始产生位置。使用此技术，可以从大量的事件中挑选来自特定区域的高能部分子喷注，它们受到温度梯度和集体流的影响基本一致，不会因为事件平均而损失信息，能够极大增强马赫锥和耗散尾流的物理信号。更详细的信息请参考文献[58]。

10 深度学习辅助Lattice QCD计算

深度学习模型可分为判别模型(分类与回归)和生成模型，生成模型能够从大量样本中学习到样本本身的统计规律，从而产生新的采样。生成模型有很多种，比如变分自编码机VAE(Variational AutoEncoder)，生成对抗网络GAN(Generative Adversarial Networks)，自回归网络，流模型Flow model以及最近出现的扩散模型Diffusion model。这些模型在图像与自然语言生成任务中的应用如雷贯耳，比如最早的视频换脸使用了VAE技术，生成以假乱真的图片使用了GAN和Flow model，最近的stable diffusion模型，可以根据文字提示生成复杂而真实的艺术图像。能够完成多轮对话，让人怀疑机

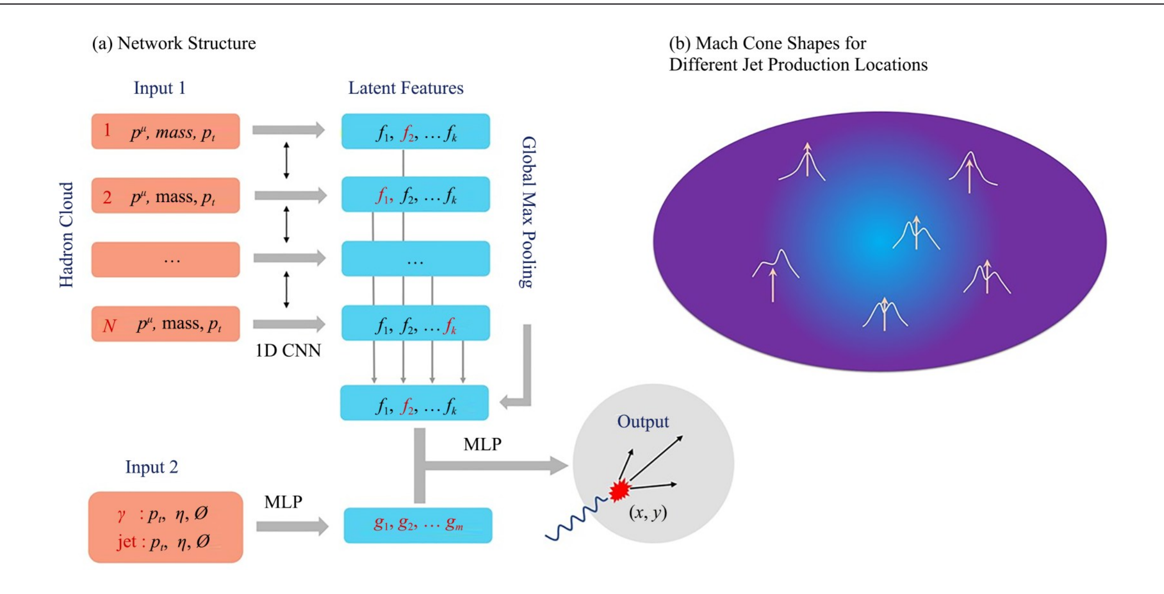


图12 深度学习辅助的喷注层析技术，定位高能部分子的初始产生位置，帮助寻找核液滴中的马赫锥[58]

Fig.12 Deep learning assisted jet tomography used to locate the initial jet production positions and aid in searching the Mach cone in nuclear liquid[58]

器已经拥有自我意识的ChatGPT则使用了自回归模型Transformer。

既然深度学习中的生成模型可以从大量样本中学习，并产生新的采样，人们自然开始思考它是否可以加速自然科学领域的蒙特卡罗采样。比如在高能核物理领域，Lattice QCD使用马尔科夫链蒙特卡罗MCMC（Markov Chain Monte Carlo）方法及其变种，比如Hybrid Monte Carlo和热浴法，从有限温度下夸克场和胶子场的配分函数采样，获得大量场的构型，来计算核物质的状态方程（即压强、能量密度、熵密度与温度的关系）。如果生成模型可以从少量的昂贵样本中学习，然后进行更大规模地采样，则可以加速第一性原理Lattice QCD的计算。文献[61]使用GAN来抽样有限温度下的1+1维复标量场。GAN包含一个生成网络，将一些随机数映射到复标量场的构型，使其尽可能地以假乱真，还包含一个判别网络，来最大化生成构型与真实构型的判别准确度。两种网络在博弈中逐渐达到均衡，此时生成网络学习到复标量场满足的分布。如果一开始的随机数满足特定的分布，比如简单的标准正态分布，则可以从标准正态分布采样一个随机数，输入经过预训练的生成模型从而产生一个复标量场的构型。图13是蒙特卡罗方法与生成对抗网络方法采样得到的复标量场平方的概率密度分布，两者相差不大。重要的是，训练好的生成模型可以并行采样，而不必像MCMC模型一样只能以序列化的方式串行采样。

MCMC方法采样的时候从一个样本出发，通过随机行走获得下一个样本，相邻样本之间存在时间上的自关联。在参数空间的临界点附近，时间上的自关联长度发散，从而使得采样效率变得极低，称作critical slowing down。文献[62−64]使用流模型和生成对抗网络辅助马尔科夫链蒙特卡罗采样，能够显著减小自关联时间，加速critical slowing down区域的采样效率。另外，训练好的流模型与GAN一样，可以使用满足简单分布的随机数，做到Lattice QCD场构型的并行采样。流模型使用了与变换法抽样相同的思路，即为了从一个复杂的概率密度分布采样，可以先从一个简单的概率密度分布采样，将采样结果进行变量替换，使得替换后的变量满足复杂的概率密度分布函数。最简单的例子是Box-Muller变换，可以从均匀分布采样，经过变换获得满足标准正态分布的样本。另一个例子是反函数法采样，求出概率密度分布的累积概率函数，0~1区间的均匀分布采样，做反函数变换，即可得到原始概率密度分布的采样。

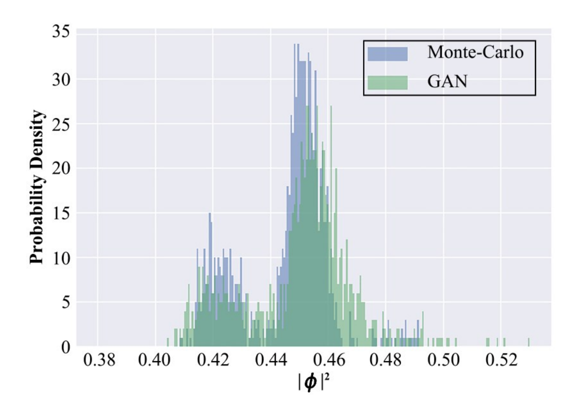


图13 蒙特卡罗方法与生成对抗网络采样得到的复标量场平方的概率密度分布对比图[61]

Fig.13 Flow chart illustrating active learning[61]

在流模型中，假设简单的概率密度分布函数为$f(z)$，复杂的概率密度分布为$p(\phi)$，可以先从$f(z)$采样得到z，使用流模型做变量替换，$\phi = g^{-1}(z)$，其中$g^{-1}: R^N \rightarrow R^N$是一个光滑、双射的流模型，根据保概率条件，这样变换得到的ϕ满足$|\tilde{p}(\phi)\mathrm{d}\phi| = |f(z)\mathrm{d}z|$，即$\phi$的分布为：

$$\tilde{p}(\phi) = f(z)\left|\frac{\mathrm{d}z}{\mathrm{d}\phi}\right| = f(z)\left|\frac{\mathrm{d}g(\phi)}{\mathrm{d}\phi}\right| \tag{27}$$

如果不加训练，则$\tilde{p}(\phi)$并不等于我们需要的概率密度分布，一般可以使用Kullback-Leibler（KL）Divergence计算两个分布之间的距离，将其定义为损失函数，通过优化来使得$\tilde{p}(\phi)$逐渐逼近$p(\phi)$。流模型是一个可逆双射的神经网络，通过构造的方法可以使得其雅克比矩阵$\left|\frac{\mathrm{d}g(\phi)}{\mathrm{d}\phi}\right|$易于计算。

最近几年，科学家发现如果把场流形通过特定的方法变换到复空间，沿着相应的路径进行积分可能消除符号问题[65–67]，但流形的选择需要专家知识以及非常复杂且昂贵的计算。机器学习被用来发现这样的积分流形，帮助消除符号问题[68–74]。如果能够完全解决符号问题，会大大推进对有限密度区域QCD相图的研究。

11 结语

本文综述了机器学习在重离子碰撞领域辅助QCD相变以及核物质状态方程研究的几种方法和思路。深度学习辅助核物质状态方程研究的方法可以分为两类，一类是直接法，试图在第一性原理计算中帮助解决Lattice QCD的采样困难以及符号问题；另一类是间接法，通过理论模型与核碰撞实验的对比，根据末态粒子谱反向推断核物质状态方程以及相变种类。这些方法包括贝叶斯推断、卷积神经网络、点云神经网、自编码机以及基于图结构的神经网络。这些网络一般被用在分类任务中，通过监督学习来建立末态粒子谱或粒子云与状态方程种类、临界涨落、喷注产生位置的非线性映射关系，或用在重建任务中，通过非监督学习提取能够反映核物质相变的序参数。这些关系都能间接帮助确定QCD相图中的核物质相变种类或临界点的位置。除此之外，深度神经网络也被用作变分函数，来表示QGP中准粒子的质量，从而能够通过自动微分和优化，学习到介质中u/d、s夸克以及胶子的有效质量，在简单的统计力学公式中重现Lattice QCD的状态方程。

致谢 感谢罗晓峰老师和马余刚老师组织这个重要的专题，感谢华为科技公司对前沿交叉学科研究的支持。

作者贡献声明 李甫鹏、庞龙刚负责文章初稿写作；庞龙刚负责文章整体设计，理论研究指导，对文章进行审阅与修订；王新年负责理论研究指导，对文章审阅与修订。

参考文献

1 Elfner H, Müller B. The exploration of hot and dense nuclear matter: introduction to relativistic heavy-ion physics[EB/OL]. 2022: arXiv: 2210.12056. https://arxiv.org/abs/2210.12056.

2 Shuryak E. Strongly coupled quark-gluon plasma in heavy ion collisions[J]. Reviews of Modern Physics, 2017, **89**(3): 035001. DOI: 10.1103/revmodphys. 89.035001.

3 Luo X F, Xu N. Search for the QCD critical point with fluctuations of conserved quantities in relativistic heavy-ion collisions at RHIC: an overview[J]. Nuclear Science and Techniques, 2017, **28**(8): 112. DOI: 10.1007/s41365-017-0257-0.

4 Bass S A, Bernhard J, Moreland J S. Determination of quark-gluon-plasma parameters from a global Bayesian analysis[J]. Nuclear Physics A, 2017, **967**: 67 – 73.

5 Everett D, Ke W, Paquet J F, *et al*. (JETSCAPE), Phenomenological constraints on the transport properties of QCD matter with data-driven model averaging[J]. Physical Review Letters, 2021, **126**: 242301.

6 Boehnlein A, Diefenthaler M, Sato N, *et al*. Colloquium: machine learning in nuclear physics[J]. Reviews of Modern Physics, 2022, **94**(3): 031003. DOI: 10.1103/revmodphys.94.031003.

7 Carleo G, Cirac I, Cranmer K, *et al*. Machine learning and the physical sciences[J]. Reviews of Modern Physics, 2019, **91**(4): 045002. DOI: 10.1103/revmodphys. 91.045002.

8 Feickert M, Nachman B. A living review of machine learning for particle physics[EB/OL]. 2021: arXiv: 2102.02770. https://arxiv.org/abs/2102.02770.

9 Das Sarma S, Deng D L, Duan L M. Machine learning meets quantum physics[J]. Physics Today, 2019, **72**(3): 48 – 54. DOI: 10.1063/PT.3.4164.

10 Mehta P, Wang C H, Day A G R, *et al*. A high-bias, low-variance introduction to Machine Learning for physicists [J]. Physics Reports, 2019, **810**: 1 – 124. DOI: 10.1016/j.physrep.2019.03.001.

11 Albertsson K, Altoe P, Anderson D, *et al*. Machine

learning in high energy physics community white paper [J]. Journal of Physics: Conference Series, 2018, **1085**: 022008. DOI: 10.1088/1742-659 6/1085/2/022008.

12 Liu J G, Wang L. Differentiate everything: a lesson from deep learning[J]. Physics, 2021, **50**(2): 69. DOI: 10.7693/wl20210201.

13 Bedaque P, Boehnlein A, Cromaz M, *et al*. A. I. for nuclear physics[J]. The European Physical Journal A, 2021, **57**(3): 100. DOI: 10.1140/epja/s10050-020-00290-x.

14 庞龙刚, 周凯, 王新年. 深度学习在核物理中的应用[J]. 原子核物理评论, 2020, **37**(3): 720 – 726. DOI: 10.11804/NuclPhysRev.37.2019CNPC41.
PANG Longgang, ZHOU Kai, WANG Xinnian. Deep learning for nuclear physics[J]. Nuclear Physics Review, 2020, **37**(3): 720 – 726. DOI: 10.11804/NuclPhysRev.37.2019CNPC41.

15 Boyle P, Bollweg D, Brower R, *et al*. Lattice QCD and the computational frontier[EB/OL]. 2022: arXiv: 2204.00039. https://arxiv.org/abs/2204.00039.

16 Monnai A, Schenke B, Shen C. Equation of state at finite densities for QCD matter in nuclear collisions[J]. Physical Review C, 2019, **100**(2): 024907. DOI: 10.1103/physrevc.100.024907.

17 Shen C, Qiu Z, Song H C, *et al*. The iEBE-VISHNU code package for relativistic heavy-ion collisions[J]. Computer Physics Communications, 2016, **199**: 61 – 85. DOI: 10.1016/j.cpc.2015.08.039.

18 Pang L G, Petersen H, Wang X N. Pseudorapidity distribution and decorrelation of anisotropic flow within the open-computing-language implementation CLVisc hydrodynamics[J]. Physical Review C, 2018, **97**(6): 064918. DOI: 10.1103/physrevc.97.064918.

19 Pratt S, Sangaline E, Sorensen P, *et al*. Constraining the equation of state of superhadronic matter from heavy-ion collisions[J]. Physical Review Letters, 2015, **114**(20): 202301. DOI: 10.1103/PhysRevLett.114.202301.

20 Bernhard J E, Moreland J S, Bass S A, *et al*. Applying Bayesian parameter estimation to relativistic heavy-ion collisions: simultaneous characterization of the initial state and quark-gluon plasma medium[J]. Physical Review C, 2016, **94**(2): 024907. DOI: 10.1103/PhysRevC.94.024907.

21 Bernhard J E. Bayesian parameter estimation for relativistic heavy-ion collisions[EB/OL]. 2018: arXiv: 1804.06469. https://arxiv.org/abs/1804.06469.

22 Bernhard J E, Moreland J S, Bass S A. Bayesian estimation of the specific shear and bulk viscosity of quark-gluon plasma[J]. Nature Physics, 2019, **15**(11): 1113 – 1117. DOI: 10.1038/s41567-019-0611-8.

23 Kuttan M O, Steinheimer J, Zhou K, *et al*. The QCD EoS of dense nuclear matter from Bayesian analysis of heavy ion collision data[EB/OL]. 2022: arXiv: 2211.11670. https://arxiv.org/abs/2211.11670.

24 Pang L G, Zhou K, Su N, *et al*. An equation-of-state-meter of quantum chromodynamics transition from deep learning[J]. Nature Communications, 2018, **9**: 210. DOI: 10.1038/s41467-017-02726-3.

25 Schenke B, Tribedy P, Venugopalan R. Fluctuating glasma initial conditions and flow in heavy ion collisions [J]. Physical Review Letters, 2012, **108**(25): 252301. DOI: 10.1103/PhysRevLett.108.252301.

26 Du Y L, Zhou K, Steinheimer J, *et al*. Identifying the nature of the QCD transition in relativistic collision of heavy nuclei with deep learning[J]. The European Physical Journal C, 2020, **80**(6): 516. DOI: 10.1140/epjc/s10052-020-8030-7.

27 Kvasiuk Y, Zabrodin E, Bravina L, *et al*. Classification of equation of state in relativistic heavy-ion collisions using deep learning[J]. Journal of High Energy Physics, 2020, **2020**(7): 133. DOI: 10.1007/JHEP07(2020)133.

28 Steinheimer J, Pang L G, Zhou K, *et al*. A machine learning study to identify spinodal clumping in high energy nuclear collisions[J]. Journal of High Energy Physics, 2019, **2019**(12): 122. DOI: 10.1007/JHEP12(2019)122.

29 Kuttan M O, Zhou K, Steinheimer J, *et al*. An equation-of-state-meter for CBM using PointNet[J]. Journal of High Energy Physics, 2021, **2021**: 184. DOI: 10.1007/JHEP10(2021)184.

30 Jiang L J, Wang L X, Zhou K. Deep learning stochastic processes with QCD phase transition[J]. Physical Review D, 2021, **103**(11): 116023. DOI: 10.1103/physrevd.103.116023.

31 Bialas A, Peschanski R B. Intermittency in multiparticle production at high energy[J]. Nuclear Physics B, 1988, **308**(4): 857 – 867. DOI: 10.1016/0550-3213(88)90131-9.

32 Satz H. Intermittency and critical behavior[J]. Nuclear Physics B, 1989, **326**(3): 613 – 618. DOI: 10.1016/0550-3213(89)90546-4.

33 Hwa R C. Fractal measures in multiparticle production[J]. Physical Review D, Particles and Fields, 1990, **41**(5):

1456－1462. DOI: 10.1103/physrevd.41.1456.

34 Antoniou N G, Contoyiannis Y F, Diakonos F K, *et al*. Pion production from a critical QCD phase[J]. Nuclear Physics A, 2001, **693**(3－4): 799－824. DOI: 10.1016/s0375-9474(01)00921-6.

35 Wu J, Lin Y F, Wu Y F, *et al*. Probing QCD critical fluctuations from intermittency analysis in relativistic heavy-ion collisions[J]. Physics Letters B, 2020, **801**: 135186. DOI: 10.1016/j.physletb.2019.135186.

36 Anticic T, Baatar B, Bartke J, *et al*. Critical fluctuations of the proton density in A+A collisions at 158A GeV[J]. The European Physical Journal C, 2015, **75**(12): 587. DOI: 10.1140/epjc/s10052-015-3738-5.

37 Davis N, Collaboration N. Searching for the critical point of strongly interacting matter in nucleus-nucleus collisions at CERN SPS[C]//Proceedings of European Physical Society Conference on High Energy Physics — PoS(EPS-HEP2019). Ghent, Belgium. Trieste, Italy: Sissa Medialab, 2020: 305. DOI: 10.22323/1.364.0305.

38 Davis N, Antoniou N, Diakonos F K, *et al*. Recent results from proton intermittency analysis in nucleus-nucleus collisions from NA61/SHINE at CERN SPS[C]//Proceedings of Corfu Summer Institute 2018 School and Workshops on Elementary Particle Physics and Gravity—PoS(CORFU2018). Corfu, Greece. Trieste, Italy: Sissa Medialab, 2019: 154. DOI: 10.22323/1.347.0154.

39 Huang Y G, Pang L G, Luo X F, *et al*. Probing criticality with deep learning in relativistic heavy-ion collisions[J]. Physics Letters B, 2022, **827**: 137001. DOI: 10.1016/j.physletb.2022.137001.

40 Bazavov A, Bhattacharya T, DeTar C, *et al*. (HotQCD), Equation of state in (2+1) -flavor QCD[J]. Physical Review D, 2014, **90**: 094503.

41 Plumari S, Alberico W M, Greco V, *et al*. Recent thermodynamic results from lattice QCD analyzed within a quasi-particle model[J]. Physical Review D, 2011, **84**: 094004.

42 Li F P, Lyu H, Pang L, *et al*. Deep-learning quasi-particle masses from QCD equation of state[EB/OL]. 2022: arXiv: 2211.07994. https://arxiv.org/abs/2211.07994.

43 Mroczek D, Hjorth-Jensen M, Noronha-Hostler J, *et al*. Mapping out the thermodynamic stability of a QCD equation of state with a critical point using active learning [EB/OL]. 2022: arXiv: 2203.13876. https://arxiv. org/abs/2203.13876.

44 Parotto P, Bluhm M, Mroczek D, *et al*. QCD equation of state matched to lattice data and exhibiting a critical point singularity[J]. Physical Review C, 2020, **101**(3): 034901. DOI: 10.1103/physrevc.101.034901.

45 Liu Z M, Zhao W B, Song H C. Principal component analysis of collective flow in relativistic heavy-ion collisions[J]. The European Physical Journal C, 2019, **79** (10): 870. DOI: 10.1140/epjc/s10052-019-7379-y.

46 Wang R, Ma Y G, Wada R, *et al*. Nuclear liquid-gas phase transition with machine learning[J]. Physical Review Research, 2020, **2**(4): 043202. DOI: 10.1103/PhysRevResearch.2.043202.

47 Liu J, Gao C, Wan N, *et al*. Basic quantities of the equation of state in isospin asymmetric nuclear matter[J]. Nuclear Science and Techniques, 2021, **32**(11): 117. DOI: 10.1007/s41365-021-00955-2.

48 Danielewicz P, Lacey R, Lynch W G. Determination of the equation of state of dense matter[J]. Science, 2002, **298**: 1592－1596.

49 Lattimer J, Prakash M. Neutron star observations: prognosis for equation of state constraints[J]. Physics Reports, 2007, **442**(1－6): 109－165. DOI: 10.1016/j.physrep.2007.02.003.

50 李甫鹏, 王永佳, 李庆峰. 利用深度学习方法研究核物质状态方程[J]. 原子核物理评论, 2020, **37**(4): 825－832. DOI: 10.11804/NuclPhysRev.37.2020017.

LI Fupeng, WANG Yongjia, LI Qingfeng. Using deep learning to study the equation of state of nuclear matter [J]. Nuclear Physics Review, 2020, **37**(4): 825－832. DOI: 10.11804/NuclPhysRev.37.2020017.

51 Liu X Q, Huang M R, Lin W P, *et al*. Symmetry energy extraction from primary fragments in intermediate heavy-ion collisions[J]. Nuclear Science and Techniques, 2015, **26**(2): S20508. DOI: 10.13538/j. 1001-8042/nst. 26. S20508.

52 Tsang M B, Zhang Y X, Danielewicz P, *et al*. Constraints on the density dependence of the symmetry energy[J]. Physical Review Letters, 2009, **102**(12): 122701. DOI: 10.1103/physrevlett.102.122701.

53 Greco V, Baran V, Colonna M, *et al*. Relativistic effects in the search for high density symmetry energy[J]. Physics Letters B, 2003, **562**: 215－220.

54 Wang Y J, Li F P, Li Q F, *et al*. Finding signatures of the nuclear symmetry energy in heavy-ion collisions with deep learning[J]. Physics Letters B, 2021, **822**: 136669. DOI: 10.1016/j.physletb.2021.136669.

55 Wang Y J, Gao Z P, Lyu H L, *et al*. Decoding the nuclear

symmetry energy event-by-event in heavy-ion collisions with machine learning[J]. Physics Letters B, 2022, **835**: 137508.

56 Ma G L, Zhang S, Ma Y G, *et al*. Di-hadron azimuthal correlation and Mach-like cone structure in a parton/hadron transport model[J]. Physics Letters B, 2006, **641**(5): 362 – 367. DOI: 10.1016/j.physletb.2006.09.001.

57 Neufeld R B. Mach cones in the quark-gluon plasma: viscosity, speed of sound, and effects of finite source structure[J]. Physical Review C, 2009, **79**(5): 054909. DOI: 10.1103/physrevc.79.054909.

58 Yang Z, He Y Y, Chen W, *et al*. Deep learning assisted jet tomography for the study of Mach cones in QGP[EB/OL]. 2022: arXiv: 2206.02393. https://arxiv. org/abs/2206.02393.

59 He Y Y, Luo T, Wang X, *et al*. Linear boltzmann transport for jet propagation in the quark-gluon plasma: elastic processes and medium recoil[J]. Physical Review C, 2015, **91**: 054908. DOI: 10.1103/PhysRevC.91.054908.

60 Cacciari M. FastJet: a code for fast k_t clustering, and more [J]. Ads Abstract Service, 2006.

61 Zhou K, Endrődi G, Pang L G, *et al*. Regressive and generative neural networks for scalar field theory[J]. Physical Review D, 2019, **100**(1): 011501. DOI: 10.1103/physrevd.100.011501.

62 Pawlowski J M, Urban J M. Reducing autocorrelation times in lattice simulations with generative adversarial networks[J]. Machine Learning: Science and Technology, 2020, **1**(4): 045011. DOI: 10.1088/2632-2153/abae73.

63 Albergo M S, Kanwar G, Shanahan P E. Flow-based generative models for Markov chain Monte Carlo in lattice field theory[J]. Physical Review D, 2019, **100**(3): 034515. DOI: 10.1103/physrevd.100.034515.

64 Wu D, Wang L, Zhang P. Solving statistical mechanics using variational autoregressive networks[J]. Physical Review Letters, 2019, **122**(8): 080602. DOI: 10.1103/PhysRevLett.122.080602.

65 Cristoforetti M, Di Renzo F, Eruzzi G, *et al*. An efficient method to compute the residual phase on a Lefschetz thimble[J]. Physical Review D, 2014, **89**(11): 114505.

66 Cristoforetti A C M, Renzo F, Scorzato L. New approach to the sign problem in quantum field theories: high density QCD on a Lefschetz thimble[J]. Physical Review D, 2012, **86**: 074506. DOI: 10.1103/PhysRevD. 86.074506.

67 Fujii H, Kamata S, Kikukawa Y. Monte Carlo study of Lefschetz thimble structure in one-dimensional Thirring model at finite density[J]. Journal of High Energy Physics, 2015, **2015**(12): 1 – 17. DOI: 10.1007/JHEP12(2015)125.

68 Alexandru A, Bedaque P, Lamm H, *et al*. Deep learning beyond Lefschetz thimbles[J]. Physical Review D, **96**(9): 094505. DOI: 10.1103/PhysRevD.96.094505.

69 Alexandru A, Bedaque P F, Lamm H, *et al*. Finite-density Monte Carlo calculations on sign-optimized manifolds[J]. Physical Review D, 2018, **97**(9): 094510. DOI: 10.1103/physrevd.97.094510.

70 Alexandru A, Bedaque P F, Lamm H, *et al*. Fermions at finite density in 2+1 dimensions with sign-optimized manifolds[J]. Physical Review Letters, 2018, **121**(19): 191602. DOI: 10.1103/PhysRevLett.121.191602.

71 Bursa F, Kroyter M. A simple approach towards the sign problem using path optimisation[J]. Journal of High Energy Physics, 2018, **2018**(12): 54. DOI: 10.1007/JHEP12(2018)054.

72 Ohnishi A, Mori Y, Kashiwa K. Path optimization for the sign problem in field theories using neural network[C]//Proceedings of the 8th International Conference on Quarks and Nuclear Physics (QNP2018). Tsukuba, Japan. Journal of the Physical Society of Japan, 2019: 024011. DOI: 10.7566/jpscp.26.024011.

73 Wynen J L, Berkowitz E, Krieg S, *et al*. Machine learning to alleviate Hubbard model sign problems[J]. Physical Review B, 2021, **103**(12): 125153. DOI: 10.1103/PhysRevB.103.125153.

74 Wan Z Q, Zhang S X, Yao H. Mitigating the fermion sign problem by automatic differentiation[J]. Physical Review B, 2022, **106**(24): L241109. DOI: 10.1103/PhysRevB.106.L241109.